

AD _____

GRANT NUMBER DAMD17-94-J-4170

TITLE: Regulation of Membrane Proteases Associated with Breast Cancer

PRINCIPAL INVESTIGATOR: Leslie Goldstein, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University
Washington, DC 20057

REPORT DATE: August 1997

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19980114 099

DTIC QUALITY INSPECTED 3

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY <i>(Leave blank)</i>	2. REPORT DATE August 1997	3. REPORT TYPE AND DATES COVERED Final (1 Aug 94 - 31 Jul 97)	
4. TITLE AND SUBTITLE Regulation of Membrane Proteases Associated with Breast Cancer		5. FUNDING NUMBERS DAMD17-94-J-4170	
6. AUTHOR(S) Leslie Goldstein, Ph.D.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Georgetown University Washington, DC 20057		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command Fort Detrick, Frederick, MD 21702-5012		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited		12b. DISTRIBUTION CODE	
13. ABSTRACT <i>(Maximum 200)</i> Seprase is a homodimeric 170 kDa integral membrane gelatinase whose expression correlates with the invasiveness of the human melanoma cell line LOX. We are interested in determining its potential role in breast cancer metastasis since seprase mRNA has been detected in the breast carcinoma cell lines MDA-MB-436 and Hs 578T. Sequence analysis of a cDNA clone from LOX indicates that it is a type II integral membrane glycoprotein that belongs to the family of nonclassical serine hydrolases. Additional sequence analysis of seprase cDNA clones from LOX, MDA-MB-436 and the fibroblast line WI-38 strongly suggests that seprase and fibroblast activation protein α are the same protein. We have produced stable seprase transfectants for the breast carcinoma cell lines MCF-7, MDA-MB-231 and MDA-MB-436 which will be tested in <i>in vitro</i> matrix degradation/invasion assay systems. Unexpectedly, we have discovered that seprase pre-mRNA undergoes alternative exon splicing giving rise to a truncated mRNA which cannot encode the membrane form of seprase but potentially encodes a truncated intracellular form of seprase. Preliminary RNA analysis of Hs 578Bst and Hs 578T suggests that the truncated RNA is up-regulated in the breast carcinoma.			
14. SUBJECT TERMS Breast Cancer Melanoma, protease, seprase, cDNA, DPPIV, FAP α		15. NUMBER OF PAGES 23	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

L.6. In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

L.6. In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

L.6. In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.


Leslie Goldstein
PI - Signature

8/29/97
Date

TABLE OF CONTENTS

Page	
1.	Cover Page
2.	SF 298
3.	Foreword
4.	Table of Contents
5-6	Introduction
7-11	Body
12	Conclusions
13-14	References
15	Bibliography
16-23	Appendix
16-17	Figure Legends
18	Figure 1
19	Figure 2
20	Figure 3
21	Figure 4
22	Figure 5
23	Figure 6

INTRODUCTION

Background

Degradation of the extracellular matrix (ECM) by proteases is a fundamental characteristic of normal tissue remodeling in physiological processes such as angiogenesis, wound healing and macrophage migration etc. However, it is well established that the pathological process of tumor cell metastases also involves the presence of several different classes of secreted and membrane bound proteases at the cell-matrix interface (1-9). It has also been shown that tumor cells elaborate extracellular ventral membrane protrusions termed “invadopodia” at the cell-matrix interface upon which certain membrane and secreted proteases are reported to be localized (10-12). Indeed, there appears to be a direct correlation between invasiveness and the elaboration of invadopodia which dissolve the surrounding matrix.(13). Furthermore, by utilizing an *in vitro* assay system that consists of growing cells on fluorescence labeled or radiolabeled ECM components that are covalently crosslinked to gelatin film, we can detect the elaboration of invadopodia by human cancer cell lines (14).

Using this *in vitro* extracellular matrix (ECM) degradation/invasion assay our laboratory identified a highly invasive melanoma cell line, LOX, that had previously been shown to have a high incidence of lung metastasis in athymic mice (15). LOX cells express a 170 kDa membrane-bound gelatinase, seprase, whose expression correlates with LOX invasiveness in the *in vitro* ECM degradation/invasion assay. Seprase expression was also detected in the breast carcinoma cell line MDA-MB 436 and more recently in the breast carcinoma line Hs 578T and the normal myoepithelial line Hs 578Bst. It is a homodimer consisting of 97 kDa subunits that are proteolytically inactive (17). Sequence analysis of 3 internal proteolytic fragments of the 97-kDa polypeptide revealed highest homology (67-88%) to the 95-kDa fibroblast activation protein α (FAP α), an inducible cell surface glycoprotein which is reported to be expressed on reactive stromal fibroblasts of epithelial cancers including breast cancer and healing wounds but whose function is unknown(18), and to a lower extent (33-70%) with the 110-kDa subunit of dipeptidyl peptidase IV (DPPIV). Dipeptidyl peptidase IV is a multifunctional protein that has wide tissue distribution and is a member of the nonclassical serine hydrolase family (19-21). Like the 170-kDa seprase its protease activity is dependent on the association of its 110 kDa subunits (22,23). We reported last year (24) the molecular cloning of a cDNA from LOX that encodes the 97-kDa subunit of seprase. Its deduced amino acid sequence confirmed the above mentioned peptide sequence data and revealed that seprase is a type II integral membrane protein with a cytoplasmic tail of 6 amino acids, followed by a transmembrane domain of 20 amino

acids and an extracellular domain of 734 amino acids. And that its carboxy terminus contains a catalytic region (~200 amino acids) which is homologous (68% identity) to that of DPPIV. We also sequenced seprase cDNA isolated from the breast carcinoma cell line MDA-MB-436 which revealed no significant difference in their primary structures. However, the seprase cDNA clones did show what appeared to be a significant sequence divergence (45 contiguous amino acids) in their catalytic regions when compared to deduced amino acid sequence predicted by the FAP α cDNA (18; see Body of Report). Otherwise, seprase and FAP α were found to be highly homologous (94% identity). We also showed that the seprase cDNA encoded a functional 170-kDa gelatinase in transiently transfected COS-7 cells.

Purpose of present work

We initially proposed to determine the possible role of seprase in the metastasis of breast cancer. The *in vitro* breast carcinoma line MDA-MB-436 which is known to express seprase would appear to be a good candidate for transfection experiments with sense and antisense cDNA constructs that encode 97-kDa subunit of seprase. Using the *in vitro* ECM invasion/degradation assay (Introduction) and also a newly developed *in vitro* 3 dimensional collagen degradation assay (see Body) we could monitor the effects of the overexpression and underexpression of this gelatinase on the invasive phenotype of MDA-MB-436 as well as its effects on other seprase negative cell lines (MCF-7, MDA-MB-231 etc.) at various stages of neoplastic development. Thus the results from these experiments might suggest and/or confirm a role for this membrane bound protease in the invasion of the ECM.

Scope of research

Our initial proposal to determine the possible role of seprase in the metastasis of breast cancer was to involve the isolation and sequence determination of seprase cDNAs, transfection studies on breast carcinoma cell lines and if possible mutagenesis studies on seprase subunit association. Importantly and unexpectedly, in the course of carrying out this grant proposal we have identified an alternatively spliced seprase mRNA transcript which potentially encodes an intracellular protease (see Body) which would therefore be distinct in function and size from the 170-kDa membrane bound form.

BODY

Experimental methods

Reverse Transcriptase - Polymerase Chain Reaction - Total RNA was isolated from tumor cells using the RNA Stat-60 kit (Tel-Test "B", Inc.). Six μ g of total RNA were used as template for first strand cDNA synthesis with oligo (dT) 12-18 as primer. The reaction was catalyzed by Superscript II Rnase H⁻ reverse transcriptase as directed by the manufacturer (Gibco/BRL). Ten oligonucleotide primers were synthesized corresponding to sense and antisense orientations based on the published FAP α cDNA sequence (18). The primers were utilized for both polymerase chain reaction (PCR) and DNA sequence analysis (see below). The 6 sense primers and their corresponding FAP α nucleotide positions were: FAP-1 5'-CCACGCTCTGAAGACAGAATT-3' (#161-181; 5' untranslated region (UTR)); FAP-3 5'-CCAGCAATGATAGCCTCAA-3' (#1055-1073); FAP-5 5'-TGACAAAATCCTCTATGCAG-3' (#1951-1971); FAP-8 5'-TCCAAGCAAGAAGTGTGTTA-3' (#1507-1526); FAP-11 5'-TTACATCTATGACCTTAGCA-3' (#598-617); FAP-13 5'-GAAACTTGGCACGGTATTCAA-3' (#45-65; 5' UTR). The 4 antisense primers and their corresponding nucleotide positions were: FAP-4 5'-ACAGACCTTACACTCTGAC-3' (#1863-1845); FAP-6 5'-TCAGATTCTGATACAGGCT-3' (#2523-2505; 3' UTR); FAP-10 5'-TAACACACTTCTGCTTGGAA-3' (#1526-1507); FAP-12 5'AACACTGTGTCAAAGCAA-3' (#2731-2713; 3' UTR). In addition, an oligonucleotide primer corresponding to the T7 promoter of the pCR3.1 vector (Invitrogen) was synthesized: 5'-TAATACGACTCACTATAGGG-3'. Polymerase chain reactions were carried out using 2 μ l of the reverse transcriptase reactions and recombinant TAQ polymerase using the manufacturers' protocol (Gibco/BRL) on a Perkin Elmer Gene Amp 9600 Cycler. One PCR cycle: Denaturation at 94 °C for 30 s; Anneal at 55 °C for 20 s and Extend at 72 °C for 30 s for a total of 40 cycles.

DNA Cloning - A ~1.2 kb amplicon generated with the FAP 1+6 primers was isolated from a 1% agarose gel using the Qiagen gel extraction kit. Purified cDNA (20ng) was ligated to 30 ng of the pCR3.1 mammalian expression vector. Ligation, transformation and selection of recombinant clones were carried out using the Eukaryotic TA Cloning Kit from Invitrogen. Purified recombinant plasmid DNA was obtained using Qiagen Tip 20 and Tip 500 columns.

DNA Sequence Analysis - Cycle sequencing reactions were carried out using the ABI PRISM Dye Terminator Ready Reaction Kit (Perkin Elmer). Automated sequencing was carried out on an ABI PRISM 377 DNA sequencer.

In Vitro Transcription - *In vitro* transcription of seprase cDNA was carried out using the T7 Cap-Scribe Kit following the manufacturers' instructions (Boehringer Mannheim). Both linearized plasmid and PCR generated amplicons using the T7 and FAP-6 or FAP-12 primers were used as templates for RNA transcription.

In Vitro Translation - *In vitro* translation was carried out using both Wheat Germ Extract and the Flexi Rabbit Reticulocyte Lysate System (Promega) following the

manufacturers' protocol. The final potassium and magnesium ion concentrations were set to 73mM and 2.1mM for the wheat germ extract translation and 86mM and 1.2mM for the rabbit reticulocyte translation. Translation products were labeled with L-[4,5-³H] leucine (152Ci/mmol) from Amersham.

Coupled Transcription -Translation - Coupled *in vitro* transcription and translation was carried out using the TnT T7 coupled reticulocyte lysate system from Promega as described. Plasmids pA11 (vector), pA15 (full length seprase; FAP 1+6), pA12 (truncated seprase; FAP 1+6) and p1425 (truncated seprase; FAP 12+13) were used at a concentration of ~1 μ g per 25 μ l reaction.

Stable Transfection of In Vitro Breast Carcinoma Cell Lines - Stable transfection of the breast carcinoma lines MCF-7, MDA-MB-436 and MDA-MB-231 were carried out with the plasmids (see above) pA11, pA15 and pA12 using the lipofectamine reagent from Gibco/BRL following the manufacturers protocol except for the following modifications:

- A) The complete growth medium was "T" medium - 1:1 RPMI + DMEM (Gibco/BRL), 10% Bovine calf serum (HyClone), 5% Nu-Serum IV (Collaborative Biomedical), 1% L-glutamine (Gibco/BRL) and 1% Penicillin and streptomycin sulfate (Gibco/BRL).
- B) The stable transfection selection medium consisted of "T" medium + G418 at concentrations of 500 μ g/ml (MCF-7), 300 μ g/ml (MDA-MB-436) and 800 μ g/ml (MDA-MB-231).
- C) Solution A - 1.5 μ g of each plasmid; Solution B - 5 μ l of Lipofectamine reagent.
- D) Solution A+B was incubated 30 minutes.
- E) The DNA-liposome complex was incubated with the cells for 6 hours.

All other procedures for isolating stable transfectant clones were as described (25).

Procedures, Results and Discussion

Seprase and FAP α are the same protein

In our midterm report (see Introduction) we stated that seprase and FAP α exhibited a significant difference in their catalytic regions that was due to the insertion of 3 guanine nucleotides in the region #1876-G to #2010-G in seprase and that FAP α because of this sequence divergence might not be a functional protease. The possibility that this divergence in sequence was due to alternative RNA processing was considered unlikely when we examined the genomic organization of the human (26) and mouse (27) DPPIV genes. We found that the corresponding genomic sequences in the DPPIV genes that would encode the region of divergence (Gly⁶²⁶ to Lys⁶⁷⁰) are contained within 2 contiguous exons (exon 22 and 23) but that the divergence in sequence does not align with the exon-intron junctions. Furthermore, the high degree sequence identity (80%) and the lack of any insertions or deletions of amino acids in both the seprase and DPPIV cDNA sequences that are encoded by exon 22 and 23 suggests that no exon-intron junction sliding has occurred between the seprase and DPPIV genes. Additionally, we sequenced the region of divergence in 13 other cDNA clones including one clone from the WI-38 fibroblast cell line. All clones exhibited the insertion of 3 guanine nucleotides that correspond with the seprase cDNA sequence. That seprase and FAP α are identical in this region was confirmed this year when the FAP α cDNA sequence was updated

(GenBank U09278). Although 3 nonconservative amino acid substitutions at positions #207 (Pro to Ala), #229 (Lys to Thr) and #354 (Arg to Thr) remain to be resolved it is clear that seprase and FAP α are the same protein.

Seprase RNA undergoes alternative splicing

In the process of cloning a cDNA by RT-PCR of LOX RNA that would encode the complete open reading frame (ORF) of seprase we detected 2 major bands (at ~ 2.4 and 1.2 kb) when using the primer pair FAP 1+6 which corresponds to sequences within the 5' UTR (FAP 1) and 3' UTR (FAP 6) of seprase mRNA (Fig. 1, lane 6, pg.18). Indeed, upon further investigation by scanning the length of the seprase mRNA by RT-PCR with primers that correspond to sequences that begin at nucleotide positions # 45 (FAP 13), # 161 (FAP 1), # 598 (FAP 11), # 1055 (FAP-3), # 1526 (FAP 10), #1863 (FAP-4) and # 2523 (FAP-6) we found that primers FAP-11 and FAP-3 only generated single bands (Fig. 1, lanes 3,5 and 7) while the other primers generated two major bands (Fig.1, lanes 1,2,4 and 6). These results suggested that a nucleotide region between #161 and # 1526 that is ~ 1.2 kb in length is deleted in some of the seprase RNA transcripts.

This finding was confirmed when we sequenced several independent clones from LOX and 2 from the breast carcinoma line MDA-MB-436. In fact, this truncated transcript has been detected (by RT-PCR) in every cell line, normal or pathological, that expresses seprase including Hs 578Bst and its malignant counterpart Hs 578T (see below). The nucleotide sequence of the truncated RNA transcript is presented in Fig. 2 (pg.19). The sequence indicates that alternative exon splicing is occurring in the seprase pre-mRNA transcript which obeys the consensus sequence motifs for splice donor and receptor sites (28). The nucleotide sequence corresponding to positions #221 to #1443 (# 13 to #1235 for the ORF; Fig. 2) is deleted which causes a frame shift in the ORF leading to the premature termination of full length seprase as a pentapeptide. Interestingly, the region of the seprase mRNA which is deleted in this truncated transcript would normally encode the cytoplasmic tail, transmembrane region, membrane proximal N-glycosylation and Cys rich regions. However, the sequence encoding the catalytic domain (# 1678 to # 2280 for the ORF) remains intact.

Analysis of the truncated mRNA downstream of the premature termination codon for full length seprase indicated that the first AUG codon that had a favorable sequence for initiation of translation (29) and that was not followed by a premature termination codon was in the seprase reading frame at positions # 1564 to # 1566 for the full length ORF. Importantly, if this downstream AUG codon was functional, the structure of the truncated mRNA would resemble a subset of mRNAs that are known to encode oncogenes, transcription factors, signal transduction components etc. that possess an upstream minicistron which is followed by a downstream coding sequence (30,31). To test the hypothesis that the truncated mRNA reinitiated protein synthesis at this downstream AUG codon we carried out *in vitro* translation with the truncated RNA (see Materials and Methods) using wheat germ extract and rabbit reticulocyte lysate. As can be seen in Fig. 3, lanes 3,4,6,7 and 8 (pg.20) both the wheat germ (lanes 3 and 4) and Flexi rabbit reticulocyte lysate (lanes 6 and 7) systems exhibit 2 bands at ~ 26 and 28 kDa. The coupled transcription –translation assay with plasmids pA12 and p1425 that encode the truncated mRNA produce a very strong band at ~ 28 kDa. The size of the *in vitro* translation products generated in this experiment are in good agreement with the predicted size of ~27 kDa for the downstream translation product. Additionally,

computer analysis of all possible ORFs in the truncated cDNA indicated that the next largest ORF would also be in the seprase reading frame and would encode a protein of ~ 22 kDa.

Using RT-PCR we have examined several *in vitro* cell lines qualitatively for the production and relative abundance of the 2 seprase mRNA transcripts. As mentioned above, the induction of the truncated transcript is not specific for either normal or malignant cell lines but is produced in all cell lines that we have examined that express seprase. However, in examining the Hs 578Bst (normal, myoepithelial; Fig. 4 (pg.21), lanes 1 and 2) and the Hs 578T (carcinoma; Fig. 4, lanes 3 and 4) *in vitro* cell lines, grown in the presence or absence of medium containing serum, we obtained a strong signal differential for the truncated transcript. It should be noted that we have carried out RT-PCR with other RNA preparations from these cell lines with less signal differential. Nevertheless, it appears that in Hs 578T there is a trend toward increased production of the truncated transcript in the presence of medium containing serum (Fig. 4, lane 3).

Stable transfection of breast carcinoma cell lines

We have isolated stable transfectants (see Materials and Methods) for the plasmids pA11 (vector), pA15 (full length seprase) and pA12 (truncated seprase) for 3 different breast carcinoma lines: MCF-7 (seprase negative, low *in vitro* invasiveness (32)), MDA-MB-436 (seprase positive, moderate *in vitro* invasiveness) and MDA-MB-231 (seprase negative, high *in vitro* invasiveness). Additionally, we have also isolated a stable transfectant for the MDA-MB-436 cell line that was transfected with pA26 (seprase cDNA in antisense orientation). We examined the transfected cell lines using RT-PCR and Western blot analyses to confirm the expression of seprase.

Using the oligonucleotide primers FAP 1+4, we detected the presence of the full length transcript in MCF-7 cells (Fig. 5, lane 3; pg.22) transfected with pA15 but it was absent in the mock (pA11) transfected MCF-7 cells (Fig. 5, lane 2). Likewise, we found the same result in MDA-MB-231 cells that had been transfected with pA15 (Fig. 5, lane 5) and pA11 (Fig. 5, lane 4). Also, the MCF-7 cells transfected with pA15 (Fig. 5, lane 3) appear to show the presence of the truncated transcript. Thus it would appear that the full length transcript generated from the pA15 cDNA that is integrated into the MCF-7 genome undergoes alternative exon splicing. RT-PCR analysis of MDA-MB-436 cells that were transfected with pA11 (Fig. 5, lane 6), pA26 (lane 7) and pA15 (lane 8) is less clear cut. This cell line exhibits the endogenous expression of seprase and thus shows 2 major bands with the FAP 1+4 primer pair in the mock transfected cells (lane 6). The pA15 transfected cell line appears to show a slight increase in the full length transcript but clearly not a marked increase in this transcript. The pA26 transfected cells (lane 7) unexpectedly appear to show a decrease in both the endogenous full length and truncated mRNAs.

We also examined the stable transfectants for the expression of the 170-kDa seprase by Western blot analysis (Fig. 6, pg.23). Seprase expression correlated with the RT-PCR results obtained for the MCF-7 pA15 transfectant (lane 5) and the MCF-7 mock transfected cells (lane 4). Likewise, MDA-MB-231 transfected with pA15 (lane 3) expresses seprase but the mock transfected cells do not (lane 2). As was the case with RT-PCR, the expression results with MDA-MB-436 is less clear cut. The mock transfected cells (lane 6) as expected express endogenous seprase; however, the pA15 transfected cells (lane 7) in agreement with the RT-PCR results exhibit either the same or

a slightly greater level of seprase but certainly not a marked increase in seprase expression. Finally, the pA26 transfected cells (lane 8) do not show any indication of decreased seprase expression. The reason for the failure of this stable transfectant to suppress endogenous seprase mRNA translation is unknown.

We now plan to test these stable transfectants in *in vitro* matrix degradation/invasion assays. Our laboratory is currently developing a new *in vitro* assay system in which the transfected cells will be encapsulated in a 3-D collagen network which more closely resembles the *in vivo* matrix environment. It will allow us to measure any change in the invasive phenotype of cells transfected with the pA15 plasmid that encodes 170-kDa membrane form of seprase. We are beginning to test the various stable transfected lines simultaneously with the submission of this report.

CONCLUSION

LOX cells express a 170 kDa membrane bound gelatinase, seprase, whose expression correlates with LOX invasiveness in the *in vitro* ECM degradation/invasion assay (16). We initially proposed to identify the 170 kDa membrane bound protease, seprase, by isolating a cDNA clone that encoded its 97 kDa subunit. And we were also interested in determining its potential role in breast cancer metastasis. As a result of our research over the period of this grant we can now definitively describe this protease at primary structure level. Its deduced amino acid sequence clearly identifies it as a member of the nonclassical serine hydrolase family. Furthermore, its high degree of identity (52%) with the exopeptidase DPPIV and the fact that both proteases require subunit association for biological activity suggests that they represent a new subfamily of serine integral membrane proteases (SIMP). Importantly, determination of the deduced amino acid sequence of seprase has clarified and confirmed its relationship to FAP α which is reported to be expressed on reactive stromal fibroblasts of various carcinomas including breast cancer but whose function was formerly unknown. We now know that seprase and FAP α are in fact the same protein. Also, the tissue expression pattern of FAP α may be different from what was formerly thought. It had not been found to be expressed on *in vitro* carcinoma cell lines (33), however, we have identified two carcinoma cell lines, MDA-MB-436 and Hs 578T, that both express seprase/FAP α . It therefore remains to be determined if the *in vivo* expression pattern of seprase/FAP α is limited to stromal fibroblasts in carcinomas. We have now produced stable seprase transfectants of breast cancer cell lines that represent different levels of neoplastic development in terms of *in vitro* invasiveness and we will now attempt to determine if expression of membrane bound seprase/FAP α changes their metastatic/invasive potential. Unexpectedly, we have made what appears to be a significant finding concerning the molecular biology of seprase/FAP α . The seprase pre-mRNA undergoes alternative exon splicing giving rise to a truncated mRNA that cannot encode the membrane form of seprase/FAP α . The existence of this truncated mRNA would appear to represent a cellular down regulation mechanism for seprase which is carried out at the level of RNA splicing. Although the induction of seprase RNA alternative processing is not tumor cell specific, preliminary results from the cell lines Hs 578Bst and Hs 578T suggest that there may be an up regulation of the truncated mRNA in the tumor cell line. Importantly, sequence analysis of the truncated mRNA indicates that a potential ORF exists downstream of the prematurely terminated seprase full length ORF. This downstream ORF would encode a truncated intracellular form of seprase (MW ~27 kDa) that would include the entire catalytic region of the full length membrane form. *In vitro* translation of the truncated mRNA confirms that the downstream ORF exists. We will now attempt to isolate the intracellular form of seprase and elucidate its role in the biology of seprase/FAP α .

REFERENCES

1. P.H.A. QuaxVan, G.N.P. Muijen, E.J.D. Weening-Verhoeff, L.R. Lund, K. Dano, D.J. Ruitterand, J.H Verheijen, *J. Cell Biol.* **115**, 191 (1991).
2. Y. Hagia, H. Fukao, S. Veshima, H. Inufusa, S. Izaki, H. Kamiishi, K. Okada and O. Matsuo, *Thromb. Res.* **65**, 449 (1992).
3. B.F. Sloane, *Semin Cancer Biol.* **1**, 137 (1990).
4. M.R. Buck, D.G. Karustis, N.A. Day, K.V. Honn, and B.F. Sloane, *Biochem. J.* **282**, 273 (1992).
5. N.S. Templeton and W.G. Stetler-Stevenson, *Cancer Res.* **51**, 6190 (1991).
6. K. Tryggvason, M. Hoyhta, C. Pyke, *Breast Cancer Res. Treat.* **24**, 209 (1993).
7. H. Sato, Y. Kida, M. Mai, Y. Endo, T. Sasaki, J. Tanaka, M. Seiki, *Oncogene* **7**, 77 (1992).
8. H. Rochefort, *Eur. J. Cancer* **28A**, 1780 (1992).
9. H. Rochefort, *Semin. Cancer Biol.* **1**, 153 (1990).
10. W.-T. Chen, *Enzyme Protein* **49**, 59 (1996).
11. H. Nakahara, M. Nomizu, S.K. Akiyama, Y. Yamada, Y. Yeh, W.-T. Chen, *J. Biol. Chem.* **271**, 27221 (1996).
12. H. Nakahara, L. Howard, E.W. Thompson, H. Sato, M. Seiki, Y. Yeh, W.-T. Chen, *Proc. Natl. Acad. Sci. USA* **94**, 7959 (1997).
13. T. Kelly, S.C. Mueller, Y. Yeh, W.-T. Chen, *J. Cell. Physiol.* **158**, 299 (1994).
14. W.-T. Chen, C.C. Lee, L. Goldstein, S. Bernier, C.Y. Lin, Y. Yeh, W.L. Monsky, T. Kelly, M. Dai, J.Y. Zhou, S.C. Mueller, *Breast Cancer Res. Treat.* **31**, 217 (1994).
15. A. Aoyama and W.-T. Chen, *Proc. Natl. Acad. Sci. USA* **87**, 8296 (1990).
16. W.L. Monsky, C.-Y. Lin, A. Aoyama, T. Kelly, S.C. Mueller, S.K. Akiyama, W.-T. Chen, *Cancer Res.* **54**, 5702 (1994).
17. M.L. Pineiro-Sanchez, L.A. Goldstein, J. Dodt, L. Howard, Y. Yeh, W.-T. Chen, *J. Biol. Chem.* **272**, 7595 (1997).
18. M.J. Scanlan, B.K. Mohan Raj, B. Calvo, P. Garin-Chesa, M.P. Sanz-Moncas, J.H. Healy, L.J. Old, W.J. Rettig, *Proc. Natl. Acad. Sci. USA* **91**, 5657 (1994).
19. D. Marguet, A.M. Bernard, I. Vivier, D. Darmoul, P. Naquet, M. Pierres, *J. Biol. Chem.* **267**, 2200 (1992).
20. F. David, A.M. Bernard, M. Pierresand, D. Marquet, *J. Biol. Chem.* **268**, 17247 (1993).
21. L. Polgar, *FEBS Lett.* **311**, 281 (1992).
22. G. Puschel, R. Mentlein, E. Heymann, *Eur. J. Biochem.* **126**, 359 (1982).
23. E.F.J. Walborg, S. Tsuchida, D.S. Weeden, M.W. Thomas, A. Barrick, K.D. McEntire, J.P. Allison, and D.C. Hixson, *Experimental Cell Research* **158**, 509 (1985).
24. L.A. Goldstein, G. Ghersi, M.L. Pineiro-Sanchez, M. Salamone, Y. Yeh, D. Flessate, W.-T. Chen, *Biochim. Biophys. Acta* **1361**, 11 (1997).

25. M. Whitt, L. Buonocore, J.K. Rose in *Current Protocols in Molecular Biology*. (Ausubel F.M. et al. ed) pp. 9.4.1.-9.9.3, John Wiley and Sons, Inc, New York, NY (1994).
26. C.A. Abbott, E.Baker, G.R. Sutherland, G.W McCaughan, *Immunogenetics* **40**, 331 (1994).
27. A.-M. Bernard, M.Marie-Genevieve, M. Pierres, D Marguet, *Biochemistry* **33**, 15204 (1994).
28. Y. Ohshima and Y. Gotoh, *J. Mol. Biol.* **195**, 247 (1987).
29. M. Kozak, *Nucl. Acids Res.* **15**, 8125 (1987).
30. M.Kozak, *Cell* **47**, 481 (1986).
31. M.Kozak, *Mamm. Genome* **7**, 563 (1996).
32. C.L. Sommers, S.W. Byers, E.W. Thompson, J.A. Torri, E.P. Gelmann, *Breast Cancer Res.Treat.* **31**, 325 (1994).
33. W.J. Rettig, P.Garin-Chesa, H.R. Beresford, H.F. Oettgen, M.R. Melamed, L.J. Old, *Proc. Natl. Acad. Sci. USA* **85**, 3110 (1988).

BIBLIOGRAPHY

Publications

1. L.A. Goldstein, G. Ghersi, M.L. Pineiro-Sanchez, M. Salamone, Y. Yeh, D. Flessate, W.-T. Chen, *Biochim. Biophys. Acta* **1361**, 11 (1997).
2. M.L. Pineiro-Sanchez, L.A. Goldstein, J. Dodt, L. Howard, Y. Yeh, W.-T. Chen, *J. Biol. Chem.* **272**, 7595 (1997).

Abstracts

1. L.A. Goldstein, G. Ghersi, M.L. Pineiro-Sanchez, M. Salamone, Y. Yeh, D. Flessate, W.-T. Chen, *Molecular Cloning of Seprase: A Serine Integral Membrane Protease from Human Melanoma*, DOD Breast Cancer Research Program: An Era of Hope, Renaissance Hotel, Washington, D.C. (1997).

Salaried Personnel

1. Leslie A. Goldstein

APPENDIX

Figure Legends

Fig. 1. RT-PCR analysis of LOX RNA using FAP α primers. RT-PCR was carried out on LOX total RNA using FAP α primers that correspond to sequences within the 5' UTR, ORF and 3'UTR of seprase/FAP α cDNA (see Materials and Methods). Lane 1 – FAP 13+4; lane 2 – FAP1+4; lane 3 - FAP11+4; lane 4 - FAP 13+10; lane 5 - FAP 11+10; lane 6 – FAP 1+6; lane 7 – FAP 3+6; lane 8 – size markers.

Fig. 2. Nucleotide and deduced amino acid sequence of seprase mRNA splice variant. Nucleotide sequence of full length seprase which encodes complete ORF is shown. Underlined sequence represents region which is deleted in the truncated splice variant. Bold nucleotides (#1561, 1564-1567) represent downstream translation initiation site (29). The putative catalytic triad of truncated seprase (Ser¹⁰³, Asp¹⁸¹, His²¹³) and its protease consensus motif are represented by bold underlined amino acids.

Fig. 3. *In vitro* translation of truncated seprase splice variant. *In vitro* translation was carried out using full length or truncated seprase RNA transcripts generated by *in vitro* transcription and plasmids encoding truncated seprase RNA (coupled transcription-translation) in wheat germ and rabbit reticulocyte systems. See Materials and Methods for details. Lane 1 – full length transcript in wheat germ system; lane 2. – minus RNA wheat germ control; lane 3 – truncated *in vitro* transcript generated from plasmid p1425; lane 4 – same as lane 3 but plasmid pA12; lane 5 – minus RNA rabbit reticulocyte control; lane 6 – truncated *in vitro* transcript generated from p1425 in rabbit reticulocyte system; lane 7 - *in vitro* translation product generated from p1425 using coupled transcription-translation rabbit reticulocyte system; lane 8 – same as lane 7 but plasmid pA12.

Fig. 4. RT-PCR analysis of Hs 578Bst and Hs 578T RNA for seprase transcripts. RT-PCR analysis was carried out on Hs 578Bst and Hs 578T total RNA using the primer pair FAP 1+4 which is capable of detecting both the full length and truncated seprase mRNAs. Note that the minor band at ~1 kb in lanes 1-4 is an artifact of PCR. Lane 1 – Hs 578Bst RNA from cells cultured in medium plus serum; lane 2 – same as lane 1 but minus serum; lane 3 – Hs 578T RNA from cells cultured in medium plus serum; lane 4 – same as lane 3 but minus serum; lane 5 – minus reverse transcriptase negative control; lane 6 – size markers.

Fig. 5. RT-PCR analysis of RNA from MCF-7, MDA-MB-231 and MDA-MB-436 stable transfectants. RT-PCR analysis was carried out on total RNA isolated from stable transfectants generated with plasmids pA11 (vector), pA15 (full length seprase) and pA26 (antisense full length seprase) using the primer pair FAP 1+4. Lane 1 – LOX RNA positive control; lane 2 – MCF-7 transfected with pA11; lane 3 – MCF-7 transfected with pA15; lane 4 – MDA-MB-231 transfected with pA11; lane 5 – MDA-

MB-231 transfected with pA15; lane 6 – MDA-MB-436 transfected with pA11; lane 7 – MDA-MB-436 transfected with pA26; lane 8 – MDA-MB- 436 transfected with pA15.

Fig. 6. Western blot analysis of MCF-7, MDA-MB-231 and MDA-MB-436 stable transfectants. Western blot analysis was carried out on the detergent lysates of stable transfectants produced with the plasmids pA11 (vector), pA15 (full length seprase) and pA26 (antisense full length seprase) using the monoclonal antibody D8 (16) which is specific for the 170 kDa dimeric form of seprase. Lane 1 – wheat germ agglutinin purified LOX extract (positive control); lane 2 – MDA-MB-231 transfected with pA11; lane 3 – MDA-MB-231 transfected with pA15; lane 4 – MCF-7 transfected with pA11; lane 5 – MCF-7 transfected with pA15; lane 6 – MDA-MB-436 transfected with pA11; lane 7 – MDA-MB-436 transfected with pA15; lane 8 – MDA-MB-436 transfected with pA26.

Figure 1

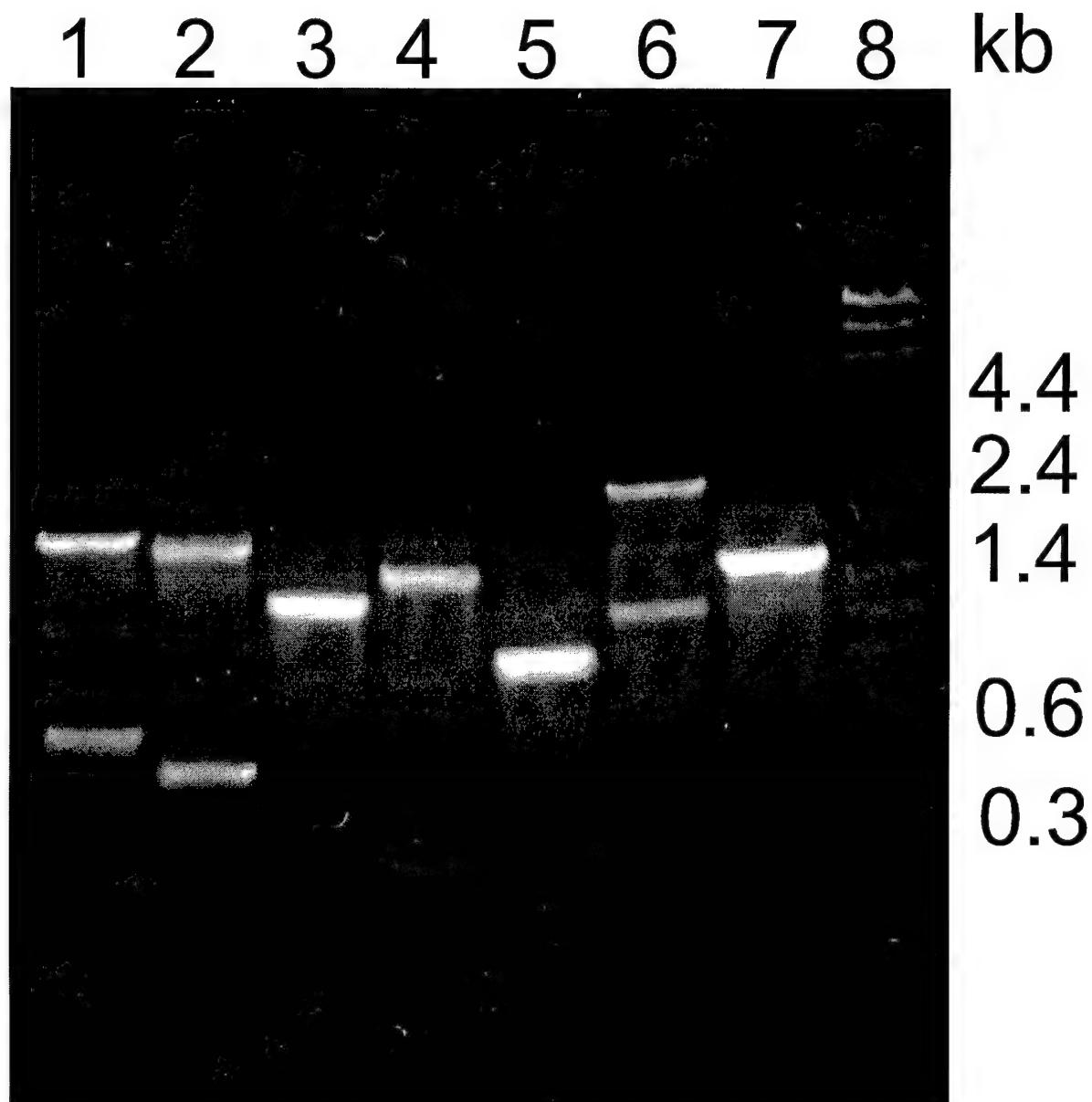


Figure 2

ATG AAG ACT TGG <u>GTA AAA ATC GTA TTT GGA GTT GCC ACC TCT GCT GTG CTT GCC TTA TTG GTG ATG TGC ATT GTC</u>	75
M K T W	4
TTA CGC CCT TCA AGA GTT CAT AAC TCT GAA GAA AAT ACA ATG AGA GCA CTC ACA CTG AAG GAT ATT TTA AAT GGA	150
<u>ACA TTT TCT TAT AAA ACA TTT TTT CCA AAC TGG ATT TCA GGA CAA GAA TAT CTT CAT CAA TCT GCA GAT AAC AAT</u>	225
<u>ATA GTA CTT TAT ATT GAA ACA GGG CAA TCA TAT ACC ATT TTG AGT AAT AGA ACC ATG AAA AGT GTG AAT GCT</u>	300
<u>TCA AAT TAC GGC TTA TCA CCT GAT CGG CAA TTT GTA TAT CTA GAA AGT GAT TAT TCA AAG CTT TGG AGA TAC TCT</u>	375
<u>TAC ACA GCA ACA TAT TAC TAT GAC CTT AGC AAT GGA GAA TTT GTA AGA GGA AAT GAG CTT CCT CGT CCA ATT</u>	450
<u>CAG TAT TTA TGC TGG CGT GGG AGT AAA TTA GCA TAT GTC TAT CAA AAC AAT ATC TAT TTG AAA CAA AGA</u>	525
<u>CCA GGA GAT CCA CCT TTT CAA ATA ACA TTT AAT GGA AGA GAA AAT AAA ATA TTT AAT GGA ATC CCA GAC TGG GTT</u>	600
<u>TAT GAA GAG GAA ATG CTT GCT ACA AAA TAT GCT CTC TGG TGG TCT CCT AAT GGA AAA TTT TTG GCA TAT GCG GAA</u>	675
<u>TTT AAT GAT ACG GAT ATA CCA GTT ATT GCC TAT TCC TAT TAT GGC GAT GAA CAA TAT CCT AGA ACA ATA AAT ATT</u>	750
<u>CCA TAC CCA AAG GCT GGA GCT AAG AAT CCC GTT CGG ATA TTT ATT ATC GAT ACC ACT TAC CCT CGC TAT GTA</u>	825
<u>GGT CCC CAG GAA GTG CCT GTT CCA GCA ATG ATA GCC TCA AGT GAT TAT TAT TTC AGT TGG CTC ACG TGG GTT ACT</u>	900
<u>GAT GAA CGA GTA TGT TTG CAG TGG CTA AAA AGA GTC CAG AAT GTT TCG GTC CTG TCT ATA TGT GAC TTC AGG GAA</u>	975
<u>GAC TGG CAG ACA TGG GAT TGT CCA AAG ACC CAG GAG CAT ATA GAA GAA AGC AGA ACT GGA TGG GCT GGT GGA TTC</u>	1050
<u>TTT GTT TCA ACA CCA GTT TTC AGC TAT GAT GCC ATT TCG TAC TAC AAA ATA TTT AGT GAC AAG GAT GGC TAC AAA</u>	1125
<u>CAT ATT CAC TAT ATC AAA GAC ACT GTG GAA AAT GCT ATT CAA ATT ACA AGT GGC AAG TGG GAG GCC ATA AAT ATA</u>	1200
<u>TTC AGA GTA ACA CAG GAT TCA CTG TTT TAT TCT AGCAA TGA A TTT GAA GAA TAC CCT GGA AGA AGA AAC ATC TAC</u>	1275

Q * 5

AGA ATT AGC ATT GGA AGC TAT CCT CCA AGC AAG AAG TGT GTT ACT TGC CAT CTA AGG AAA GAA AGG TGC CAA TAT	1350
TAC ACA GCA AGT TTC AGC GAC TAC GCC AAG TAC TAT GCA CTT GTC TGC TAC GGC CCA GGC ATC CCC ATT TCC ACC	1425
CTT CAT GAT GGA CGC ACT GAT CAA GAA ATT AAA ATC CTG GAA AAC AAG GAA TTG GAA AAT GCT TTG AAA AAT	1500
ATC CAG CTG CCT AAA GAG GAA ATT AAG AAA CTT GAA GAT GAA ATT ACT TTA TGG TAC <u>AAG ATG ATT CTT CCT</u>	1575

M I L P 4

CCT CAA TTT GAC AGA TCA AAG AAG TAT CCC TTG CTA ATT CAA GTG TAT GGT GGT CCC TGC AGT CAG AGT GTA AGG	1650
P Q F D R S K K Y P L L I Q V Y G G P C S Q S V R 29	
TCT GTA TTT GCT GTT AAT TGG ATA TCT TAT CTT GCA AGT AAG GAA GGG ATG GTC ATT GCC TTG GTG GAT GGT CGA	1725
S V F A V N W I S Y L A S K E G M V I A L V D G R 54	
GGA ACA GCT TTC CAA GGT GAC AAA CTC CTC TAT GCA GTG TAT CGA AAG CTG GGT GTT TAT GAA GTT GAA GAC CAG	1800
G T A F Q G D K L L Y A V Y R K L G V Y E V E D Q 79	
ATT ACA GCT GTC AGA AAA TTC ATA GAA ATG GGT TTC ATT GAT GAA AAA AGA ATA GCC ATA TGG GGC TGG TCC TAT	1875
I T A V R K F I E M G F I D E K R I A I W G W S Y 104	
GGA GGA TAC GTT TCA TCA CTG GCC CTT GCA TCT GGA ACT GGT CTT TTC AAA TGT GGT ATA GCA GTG GCT CCA GTC	1950
<u>G G Y V S S L A L A S G T G L F K C G I A V A P V</u> 129	
TCC AGC TGG GAA TAT TAC GCG TCT GTC TAC ACA GAG AGA TTC ATG GGT CTC CCA ACA AAG GAT GAT AAT CTT GAG	2025
S S W E Y Y A S V Y T E R F M G L P T K D D N L E 154	
CAC TAT AAG AAT TCA ACT GTG ATG GCA AGA GCA GAA TAT TTC AGA AAT GTA GAC TAT CTT CTC ATC CAC GGA ACA	2100
H Y K N S T V M A R A E Y F R N V D Y L L I H G T 179	
GCA GAT GAT AAT GTG CAC TTT CAG AAC TCA GCA CAG ATT GCT AAA GCT CTG GTT AAT GCA CAA GTG GAT TTC CAG	2175
A <u>D</u> D N V H F Q N S A Q I A K A L V N A Q V D F Q 204	
GCA ATG TGG TAC TCT GAC CAG AAC CAC GGC TTA TCC GGC CTG TCC ACG AAC CAC TTA TAC ACC CAC ATG ACC CAC	2250
A M W Y S D Q N <u>H</u> G L S G L S T N H L Y T H M T H 229	
TTC CTA AAG CAG TGT TTC TCT TTG TCA GAC TAA	2283
F L K Q C F S L S D	239

Figure 3

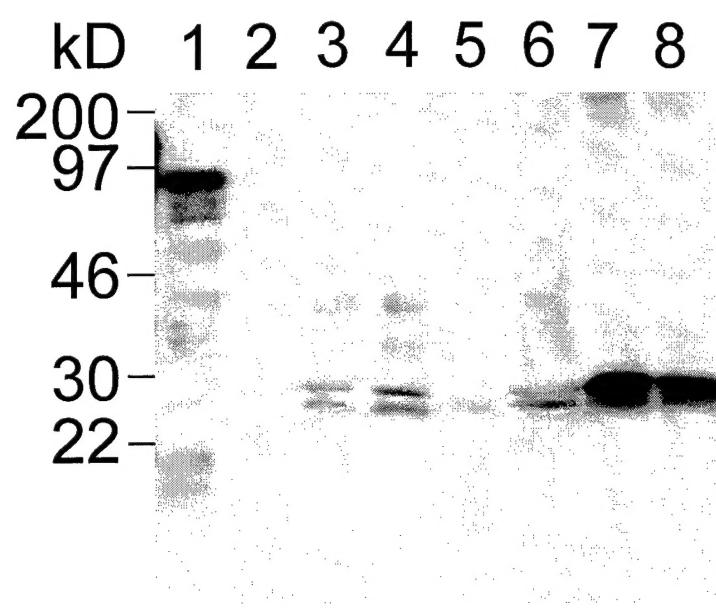


Figure 4

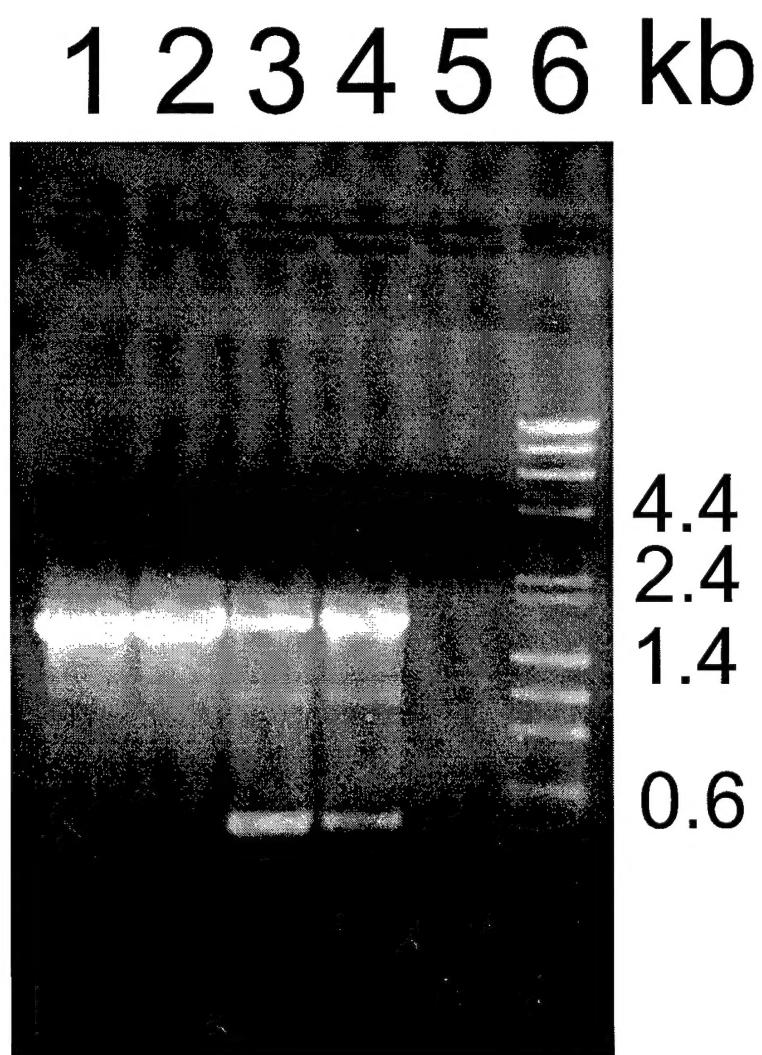


Figure 5

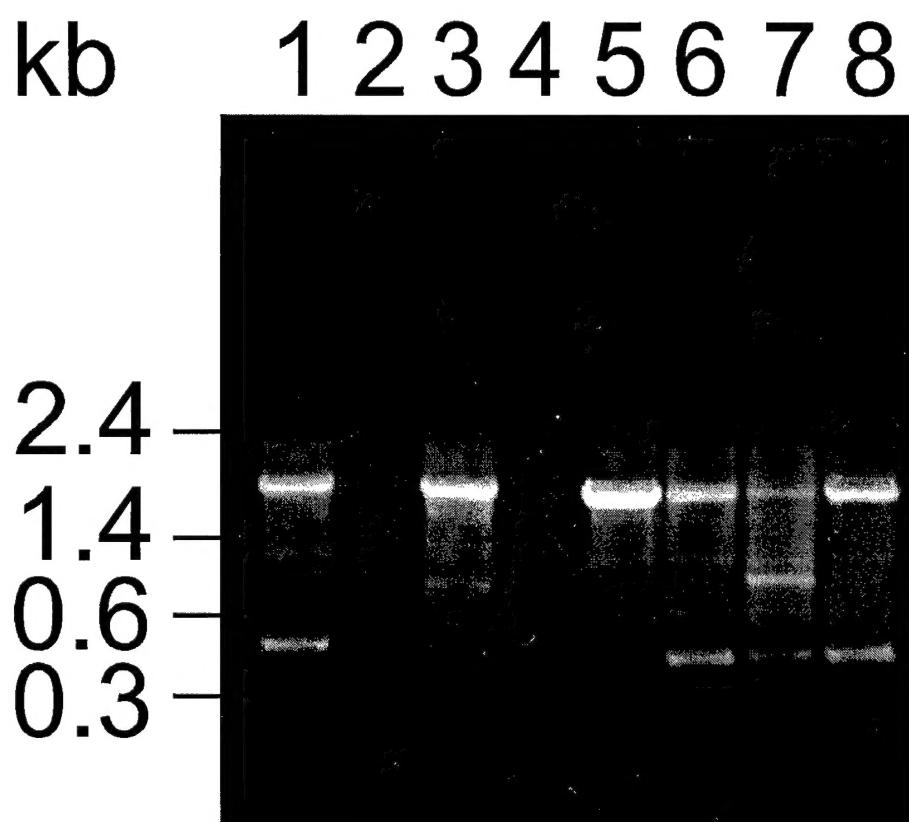


Figure 6

